The perimeter of large planar Voronoi cells: a double-stranded random walk

H.J. Hilhorst

Laboratoire de Physique Théorique Bâtiment 210, Université de Paris-Sud 91405 Orsay Cedex, France

February 2, 2008

Abstract

Let p_n be the probability for a planar Poisson-Voronoi cell to have exactly n sides. We construct the asymptotic expansion of $\log p_n$ up to terms that vanish as $n \to \infty$. Along with it comes a nearly complete understanding of the structure of the large cell. We show that two independent biased random walks executed by the polar angle determine the trajectory of the cell perimeter. We find the limit distributions of (i) the angle between two successive vertex vectors, and (ii) the one between two successive perimeter segments. We obtain the probability law for the perimeter's long wavelength deviations from circularity. We prove Lewis' law and show that it has coefficient 1/4.

1. Introduction

Cellular structures are observed in a great diversity of situations in science and technology. A common phenomenological description of such structures makes use of Voronoi cells [1], constructed around centers which for simplicity are referred to as "particles": the Voronoi cell of a particle a is made up of those points in space for which a is the closest particle. Hence Voronoi cells constitute a "tessellation", i.e. a division of space into nonoverlapping regions. An excellent review of tessellations, with in its introduction many historical and anecdotal details, is due to Okabe et al. [2]. It offers a thorough discussion of much fundamental and applied work in this area. This includes applications to biological tissues, mesh generation in numerical algorithms, coding in telecommunication, etc.

Whereas the Wigner-Seitz cell associated with the regular lattices of solid state physics is a special case, the generic Voronoi tessellation is based on randomly distributed particles; for uniformly distributed point-particles the result is called a *Poisson-Voronoi* tessellation.

Voronoi cells appear in physics for many different reasons. Meijering [3] used them to characterize the end result of a crystal growth process starting from independent centers. In the theory of the solid-liquid transition they serve to define nearest-neighbor relations between atoms, which in turn allow the identification of lattice defects. Recently Voronoi tessellations of fractal sets have been considered [4]. In field theory the use of random lattices and the associated Voronoi tessellation [5, 6] was motivated by their advantage of being statistically invariant under arbitrary translations and rotations while still preserving a short-distance cutoff.

Here we are interested in two-dimensional Poisson-Voronoi tessellations. Planar Voronoi cells are convex polygons. The first and foremost quantity that appears in both theoretical and applied work is the probability p_n for a planar Voronoi cell to be n-sided. Approximate values for p_n have been tabulated by many authors [2]; they peak at n=6 and fall off rapidly for larger n. Connected questions concern the statistics of the cell area, its perimeter, and its angles, as well as correlations with neighboring cells. They have been studied by methods ranging from heuristic arguments and Monte Carlo simulations to exact methods. Whereas some statistical properties are readily obtained, others, among which the fraction p_n , still defy solution [6, 2].

Interest in planar Voronoi cells having a large number n of sides first arose in biology. Lewis' empirical law [7, 2], formulated more than seventy years ago, states that when n becomes large, the average area of the n-sided cell grows as $\simeq c(n-2)/\nu$, where ν is the particle density. In spite of attempts, this asymptotic proportionality with n has not so far been proved [9]. Numerical estimates of the coefficient c are in the range 0.199-0.257 [8, 10, 2]. In metallography large cells were observed by Aboav [11] in the arrangement of grains in a polycrystal. His empirical law states that the neighbor of an n-sided cell has itself on average $a_0 + a_1 n^{-1}$ neighbors. Only recently Hug $et\ al.$ [12] showed rigorously that when a Voronoi cell becomes large (in an appropriate sense), its shape tends towards a circle. In a different approach mathematicians study those large cells that allow for an inscribed circle whose given radius becomes asymptotically large [13].

Drouffe and Itzykson (DI) [8] estimated the probability p_n for a planar Voronoi cell to be n-sided by Monte Carlo simulations for n up to 50. The rejection rate limits the statistical precision of generating such very rare events. All other simulations cited in Ref. [2] (see e.g. [10]) were restricted to $n \lesssim 14$, so that the work by DI is still today the main reference for the large n behavior.

This note is the first announcement of a collection of new analytical results

on planar Poisson-Voronoi tessellations. All of them are by-products of our asymptotic evaluation of p_n , and together they lead to a nearly complete understanding of the large n-sided cell. Our analysis uses an accumulation of classical mathematical methods; however, the asymptotics is beset by a great number of difficulties, of which the correct choice of variables of integration is merely the first one. It will be possible here to indicate only the main steps and the main results.

We begin by a precise statement of the question. Let a subdomain of the plane, of area L^2 , contain N particles with tags a, placed at positions $\vec{\Re}_a$ chosen independently and with uniform probability, and let $N, L \to \infty$ at fixed $N/L^2 = \nu$. We consider the Voronoi cell of an arbitrarily chosen particle a_0 whose position will be the origin of the coordinate system. The perpendicular bisectors of the $\vec{\Re}_a$ (with $a \neq a_0$) pass through the midpoints $\vec{R}_a = \frac{1}{2}\vec{\Re}_a$, which are uniformly distributed with density 4ν . Each side of the cell is a segment of one of the perpendicular bisectors. An n-sided cell may be fully specified by n midpoint vectors \vec{R}_m ($m = 1, \ldots, n$); or alternatively by the n vertex vectors \vec{S}_m ($m = 1, \ldots, n$) that connect the origin to its vertices.

The probability p_n has been expressed [8, 6, 14] as the 2n-dimensional integral

$$p_n = \frac{1}{n!} \int \left[\prod_{m=1}^n d\vec{R}_m \right] \chi e^{-A}, \qquad (1)$$

where we scaled lengths such that $4\nu = 1$; it is analogous to a partition function in statistical mechanics. The functions χ and \mathcal{A} enforce two conditions necessary for the n vectors \vec{R}_m to define a valid Voronoi cell. First, each of the n associated bisectors must contribute a nonzero segment to the perimeter. Calka [14] formulates this condition as a set of n inequalities, and χ is the indicator function of the phase space domain where they are satisfied. Secondly, the union of the n disks centered at $\frac{1}{2}\vec{S}_m$ and of radius $\frac{1}{2}S_m$ must be void of midpoints. The probability for this is $e^{-\mathcal{A}}$, where \mathcal{A} is the area of the void region. Explicit expressions for \mathcal{A} were given by DI [8] and by Calka [14]. Eq. (1) has been rewritten and evaluated to seven decimal places [15] for n = 3, but has otherwise remained untractable analytically.

2. The perimeter as a Markov process

Our approach originated from an attempt to describe the perimeter of a Voronoi cell as a Markov process with the polar angle as independent variable. We order the midpoints $\vec{R}_m \equiv (R_m, \Phi_m)$ such that $0 < \Phi_1 < \ldots < \Phi_{n-1} < \Phi_n \equiv 2\pi$ and denote by $\vec{S}_m \equiv (S_m, \Psi_m)$ the vertex where the bisectors through \vec{R}_{m-1} and \vec{R}_m intersect. Let β_m be the angle between \vec{R}_{m-1} and \vec{S}_m , and γ_m the one between \vec{S}_m and \vec{R}_m . These angles may be positive or negative, since a midpoint need not (and for large n generically will not!) itself be part of the side that it defines. The Markov process description

leads to a transition probability between two successive midpoints which has, essentially, the form $R_m^2 T_m e^{-A_m}$. Here A_m , which satisfies $A = \sum_{m=1}^n A_m$, is the area of a region delimited by three circular arcs that should be void of midpoints if the perimeter is to travel uninterruptedly between \vec{S}_{m-1} and \vec{S}_m ; and $T_m = \cos \gamma_m \sin(\beta_m + \gamma_m)/\cos^3 \beta_m$ is the (unnormalized) probability density for the perimeter to turn by $\beta_m + \gamma_m$ at vertex \vec{S}_m . Considerable algebra then gives

$$p_{n} = \frac{1}{n} \int d\beta d\gamma \, \delta\left(\sum_{m=1}^{n} (\beta_{m} + \gamma_{m}) - 2\pi\right) \delta(G)$$

$$\times \left[\prod_{m=1}^{n} \rho_{m}^{2} T_{m}\right] \int_{0}^{\infty} dR \, R^{2n-1} e^{-\pi R^{2} A(\beta, \gamma)}, \qquad (2)$$

which calls for several explanations. (i) We abbreviated

$$\int d\beta d\gamma = \int_{-\pi/2}^{\pi/2} d\beta_1 \int_{-\beta_1}^{\pi/2} d\gamma_1 \int_{-\gamma_1}^{\pi/2} d\beta_2 \int_{-\beta_2}^{\pi/2} d\gamma_2 \dots$$

$$\dots \int_{-\gamma_{n-1}}^{\pi/2} d\beta_n \int_{-\beta_n}^{\pi/2} d\gamma_n \, \theta(\gamma_n + \beta_1) \,. \tag{3}$$

Because of how they are nested, these integrations impose an orientation on the perimeter. (ii) For an arbitrary set of angles $(\beta, \gamma) \equiv \{\beta_m, \gamma_m\}$ the perimeter will spiral instead of close onto itself after a full turn of 2π ; the factor $\delta(G)$, where $G = (2\pi)^{-1} \sum_{m=1}^{n} \log(\cos \gamma_m / \cos \beta_m)$, enforces its closure. (iii) The ratios $\rho_m \equiv R_m/R$, where R is the average midpoint distance $R = n^{-1} \sum_{m=1}^{n} R_m$, are functions of the angles defined by

$$\rho_m = \frac{\cos \gamma_m \cos \gamma_{m-1} \dots \cos \gamma_1}{\cos \beta_m \cos \beta_{m-1} \dots \cos \beta_1} \rho_n \tag{4}$$

for m = 1, ..., n - 1, together with the sum rule $n^{-1} \sum_{m=1}^{n} \rho_m = 1$. (iv) Finally, we set $\mathcal{A} = \pi R^2 A(\beta, \gamma)$.

Eq. (2) fully defines our starting point. As a check one may derive it directly from (1) by a coordinate transformation, in which the factors $\rho_m^2 T_m$ then appear as the Jacobian. The decisive advantage of (2) over (1) is that the integration limits have been made explicit so that the function χ is no longer needed.

Integrating on R in (2) yields $\frac{1}{2}(n-1)! [\pi A(\beta, \gamma)]^{-n}$. If, as will appear, $A(\beta, \gamma)$ differs negligibly from unity for $n \to \infty$, a steepest descent analysis of the same integral shows that it draws its main contribution from R within a width of order n^0 around the saddle point $R_c = \sqrt{n/\pi}$.

Let now $P_n(\Omega)$ be expression (2) but with the 2π in the delta function replaced by a new variable Ω , and let $\tilde{P}_n(s)$ be its Laplace transform with Laplace variable ns. Here we pass to a "canonical ensemble"

of angles with weight $\exp[-ns\sum_{m=1}^{n}(\beta_m+\gamma_m)]$. We define $\mathbb{H}(s)$ by writing $\tilde{P}_n(s)=\int \mathrm{d}\beta\mathrm{d}\gamma \ \mathrm{e}^{-\mathbb{H}(s)}$. The evaluation of this integral constitutes the real problem: it is on all possible *shapes* (β,γ) of a Voronoi cell with average midpoint distance scaled to unity.

3. Asymptotic expansion

Anticipating that we will be able to identify a zeroth order problem we write $\mathbb{H}(s) = \mathbb{H}_0(s) + \mathbb{V}$. Hence

$$\int d\beta d\gamma \ e^{-\mathbb{H}(s)} = \langle e^{-\mathbb{V}} \rangle_0 \int d\beta d\gamma \ e^{-\mathbb{H}_0(s)}, \qquad (5)$$

where $\langle ... \rangle_0$ is the average with weight $\exp[-\mathbb{H}_0(s)]$. The appropriate choice of \mathbb{H}_0 , if there is one at all, depends delicately on how the variables of integration in (5) are assumed to scale with n. We certainly expect that for large n only small β_m and γ_m will contribute, but the exact scaling of these angles with n is not a priori evident. The calculation bears out that

$$e^{-\mathbb{H}_0(s)} = \pi^{-n} \prod_{m=1}^n (\beta_m + \gamma_m) \exp[-ns(\beta_m + \gamma_m)]$$
 (6)

is the right choice for \mathbb{H}_0 .

We now transform to another set of angles, viz. $\xi_m = \beta_m + \gamma_m$ and $\eta_m = \gamma_m + \beta_{m+1}$ where m = 1, ..., n and $\beta_{n+1} \equiv \beta_1$. Clearly ξ_m is the angle between two successive midpoint vectors and η_m the one between two successive vertex vectors. Since the set $\{\xi_m, \eta_m\}$ does not fix the relative angle between the systems of midpoint and of vertex vectors, we complete it by one of the original angles, β_1 . Setting $\Phi_m = \sum_{\ell=1}^m \xi_\ell$ and $\Psi_m = \beta_1 + \sum_{\ell=1}^{m-1} \eta_\ell$ we obtain for m = 1, ..., n the inverse relation

$$\beta_m = \Psi_m - \Phi_{m-1}, \quad \gamma_m = \Phi_m - \Psi_m. \tag{7}$$

The "zeroth order" problem is constituted by the integral that is left in the RHS of (5) when one sets $\mathbb{V}=0$. It can be solved when reformulated in terms of the ξ_m and η_m . Since ξ_m , $\eta_m > 0$ and $\sum_m \xi_m = \sum_m \eta_m = 2\pi$, they must scale as $\xi_m = n^{-1}x_m$ and $\eta_m = n^{-1}y_m$. In the limit $n \to \infty$ the x_m and y_m may be integrated from 0 to ∞ and the problem of the nested integrations disappears. One finds $\int d\beta d\gamma \exp[-\mathbb{H}_0(s)] = (3n - 2)! \pi^{-n}(ns)^{-3n+1}/(2n)!$, up to contributions that vanish exponentially with n. Let $P_n^{(0)}(\Omega)$ be the Laplace inverse of this zeroth order expression; one readily finds $P_n^{(0)}(2\pi) = (8\pi^2)^n/[4\pi^2(2n)!]$. The inverse Laplace integral has a saddle point at $s = s_c = 3/\Omega$; if we can show that for $n \to \infty$ the prefactor $\langle \exp(-\mathbb{V}) \rangle_0$ in Eq. (5) has a finite nonzero value in $s = s_c$, then it follows that $p_n = \langle \exp(-\mathbb{V}) \rangle_{3/2\pi} P_n^{(0)}(2\pi)$.

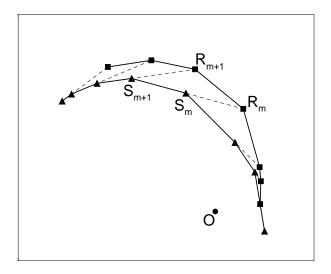


Figure 1: Schematic picture of the two random walks associated with the perimeter of a Voronoi cell. The walk joining the vertices \vec{S}_m is the actual perimeter; the other walk joins the midpoints \vec{R}_m . In the limit $n \to \infty$ the increments of the polar angle along each walk are independent random variables distributed according to Eq. (8). The perimeter segments are on the perpendicular bisectors (dashed lines).

We pause for a few comments. The zeroth order calculation has a useful corollary: in the limit $n \to \infty$ the scaled angles x_m and y_m become independent random variables with laws $u(x_m)$ and $v(y_m)$ given by

$$u(x) = \pi^{-2} x e^{-x/\pi}, \qquad v(y) = (2\pi)^{-1} e^{-y/(2\pi)}.$$
 (8)

It follows, again for $n \to \infty$, that $\{\Phi_m\}_{m=0}^n$ and $\{\Psi_m\}_{m=0}^n$ are two mutually independent one-dimensional random walks, each with a drift of $2\pi/n$ per step and with independent increments; they describe how the polar coordinate of the midpoint vector $\vec{R}_m = (R_m, \Phi_m)$ and of the vertex vector $\vec{S}_m = (S_m, \Psi_m)$, respectively, increases as one progresses along the perimeter. These angular random walks *drive* the radial coordinates R_m and S_m and couple them together, which leads to the double-stranded *two*-dimensional walk depicted in Fig. 1.

Let us now pursue the analysis. The validity of the zeroth order results hinges on our ability to prove that the first factor in Eq. (5), $\langle \exp(-\mathbb{V}) \rangle_0$, exists and is subdominant in the large n expansion of $\log p_n$. Actually, everything has been so prearranged that this factor is of order n^0 . Moreover, it has an interesting structure.

The regular n-sided polygon is an obvious point of symmetry in phase space. DI obtained an exact lower bound for p_n by expanding the vertex vectors \vec{S}_m about their regular polygon positions. This procedure preserves,

however, the long range polygonal order. Here we will advance as in the theory of elasticity, where one considers the deviations not of the positions, but of the interatomic distances, from their ground state values. Hence we expand \mathbb{V} in powers of $\delta x_m = x_m - \bar{x}$ and $\delta y_m = y_m - \bar{y}$ (where $\bar{x} = \bar{y} = 2\pi$). This procedure allows, in principle, for large deviations from the regular n-sided polygon.

Terms from various sources contribute to \mathbb{V} and have to be examined one by one. The scaling of the ξ_m and η_m and their asymptotic independence determine how the other variables scale; Eq. (7) implies, in particular, that the original angles scale as $\beta_m = n^{-\frac{1}{2}}b_m$ and $\gamma_m = n^{-\frac{1}{2}}c_m$. One finds that $\mathbb{V} = \mathbb{V}_0 + n^{-\frac{1}{2}}\mathbb{V}_1 + \ldots$ Hence $\langle \exp(-\mathbb{V}) \rangle_0 = \langle \exp(-\mathbb{V}_0) \rangle_0 [1 + \mathcal{O}(n^{-\frac{1}{2}})]$. It appears that \mathbb{V}_0 is quadratic in the δx_m and δy_m with contributions coming from (i) the $\mathcal{O}(n^{-1})$ terms in the expansion $A = 1 + n^{-1}A_1 + \ldots$, and (ii) the product $\prod_{m=1}^n \rho_m^2 T_m/(\beta_m + \gamma_m)$. In terms of the Fourier components $\hat{z}_q = (2\pi n^{\frac{1}{2}})^{-1} \sum_{m=1}^n \mathrm{e}^{2\pi \mathrm{i} q m/n} \delta z_m$, where z = x, y and q is integer, one finally finds, with $s = 3/(2\pi)$,

$$\mathbb{V}_0 = \sum_{q \neq 0} \left[(q^{-2} + 2q^{-4})(\hat{x}_q - \hat{y}_q)(\hat{x}_{-q} - \hat{y}_{-q}) + 2q^{-2}(\hat{x}_q - \hat{y}_q)\hat{y}_{-q} \right], \quad (9)$$

where q runs through all nonzero integers. The decay of the Fourier coefficients as $\sim q^{-2}$ for large q is characteristic of interacting Coulomb charges. Since the \hat{x}_q and \hat{y}_q are sums of independent variables, they are Gaussian distributed with variances determined by (8). We can show that the quantity $\exp(-\mathbb{V}_0)$ may be averaged by Gaussian integration: in terms of Coulomb charges this amounts to a Debye-Hückel approximation, which becomes exact in the high temperature limit; here the inverse "temperature" is 1/n and the exactness of our procedure follows [16] from a Hubbard-Stratonovich transformation. Gaussian integration then leads to $\langle \exp(-\mathbb{V}_0) \rangle_{3/2\pi} = C$ where $C = \prod_{q=1}^{\infty} (1-q^{-2}+4q^{-4})^{-1} = 0.344.$. Hence the finite size correction factor C is the partition function of the deformational modes of the perimeter with wavelength $2\pi/q$, the main contributions coming from the long wavelengths.

4. Results

Upon combining the zeroth and first order result we conclude that the probability p_n for a Voronoi cell to be n-sided behaves as

$$p_n = \frac{C}{4\pi^2} \frac{(8\pi^2)^n}{(2n)!} \left[1 + \mathcal{O}(n^{-\frac{1}{2}}) \right], \qquad n \to \infty,$$
 (10)

which may also be written as the asymptotic equality

$$\log p_n \simeq -2n \log n + n \log(2\pi^2 e^2) - \frac{1}{2} \log(2^6 \pi^5 C^{-2} n). \tag{11}$$

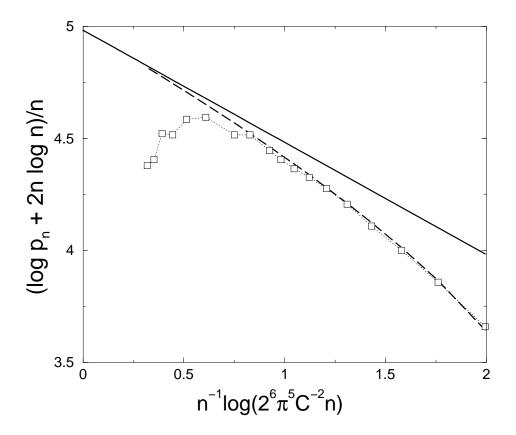


Figure 2: Straight solid line: the asymptotic formula, Eq. (11). Open squares: Monte Carlo data by Drouffe and Itzykson (Ref. [8], Table 1). Dashed line: Eq. (11) with an additional term $an^{-\frac{1}{2}}$, where we adjusted a=-0.24 to fit the data.

We compare this to existing work. DI assume an expansion similar to (11) and provide strong analytic an numerical evidence for the coefficient of the $n \log n$ term to be equal to -2. This has now been confirmed. Furthermore, DI prove that the term proportional to n has a coefficient larger than $\log(\pi^2/e) = 1.289...$, and from their numerical data estimate that it is actually between 4 and 5. Our value $\log(2\pi^2e^2) = 4.982...$ satisfies DI's exact inequality, as of course it had to, and is just inside the range of values that these authors judged likely.

In Fig. 2 our asymptotic result Eq. (11) is shown as a straight solid line of slope $-\frac{1}{2}$. It is compared to the Monte Carlo data of Ref. [8] (open squares) in the range $50 \ge n \ge 7$ (the abscissa value of 2 corresponds very nearly to n=7). The first term not displayed in the asymptotic series of Eq. (11) will be of the form $an^{-\frac{1}{2}}$; if, venturing at this point beyond our analytic results, we choose its amplitude to fit the data, very good agreement (the dashed

line) with the Monte Carlo results is obtained for $a = -0.24 \pm 0.02$. We believe that the downward trend in the DI data for $n \gtrsim 30$ (abscissa $\lesssim 0.75$) is due to the great difficulty of such simulations; in that region the authors estimated their error bars to be at least of the order of the data themselves.

Consistency of the asymptotic scaling with Eq. (4) requires that $\rho_m = 1 + n^{-\frac{1}{2}} r_m$, where r_m remains of order n^0 as $n \to \infty$. Hence the midpoint distances $R_m = R \rho_m$ deviate by order n^0 from their average R, which itself is typically within n^0 from R_c . Therefore the area of the n-sided Voronoi cell is sharply peaked around an average equal to $\pi R_c^2 = n/(4\nu)$, where we restored the particle density. This demonstrates that to leading order Lewis' law holds. Its coefficient $c = \frac{1}{4}$ agrees with DI's observation that $c \approx \frac{1}{4}$.

It is amazing that the analysis of a problem defined on only a random set of points in the plane evokes associations with so many phenomena in statistical physics. A full account [16] of this work is in preparation. In an extension of it we will revisit Aboav's law.

Acknowledgments. The author thanks Pierre Calka for a discussion and for pointing out some of the mathematical literature on this subject, and Bernard Jancovici for conversations on Coulomb charges. He thanks Jean-Michel Drouffe for correspondence. A casual remark by Emmanuel Trizac sparked this work off. The Laboratoire de Physique Théorique of the Université de Paris-Sud is associated with the Centre National de la Recherche Scientifique as research unit UMR 8627.

References

- [1] G. Voronoi, J. Reine Angew. Math. 134, 198 (1908).
- [2] A. Okabe, B. Boots, K. Sugihara, and S. N. Chiu, *Spatial tessellations: concepts and applications of Voronoi diagrams*, second edition (John Wiley & Sons Ltd., Cichester, 2000).
- [3] J. L. Meijering, Philips Research Reports 8, 270 (1953).
- [4] H. Hinrichsen and G. Schliecker, J. Phys. A 31, L451 (1998).
- [5] N. H. Christ, R. Friedberg, and T. D. Lee, Nuclear Physics B202, 89 (1982);
 B210, 310; 337 (1982).
- [6] C. Itzykson and J.M. Drouffe, *Statistical field theory* (Cambridge University Press, Cambridge, 1989), Vol. 2, chapter 11.
- [7] F. T. Lewis, Anatomical Record 38, 341 (1928); 47, 59 (1930); 50, 235 (1931).
- [8] J. M. Drouffe and C. Itzykson, Nuclear Physics **B235** [FS11], 45 (1984).

- [9] N. Rivier and A. Lissowski, J. Phys. A 15, L143 (1982); N. Rivier, Phil. Mag. B 52, 795 (1985); S. N. Chiu, J. Phys. A 28, 607 (1995).
- [10] M. P. Quine and D. F. Watson, J. Appl. Probab. 21, 548 (1984); G. Le Caër and J. S. Ho, J. Phys. A 23, 3297 (1990); S. Kumar and S. K. Kurtz, Materials Characterization 31, 55 (1993).
- [11] D.A. Aboav, Metallography 3, 383 (1970).
- [12] D. Hug, M. Reitzner, and R. Schneider, Adv. in Appl. Probab. 36, 667 (2004).
- [13] P. Calka, Adv. in Appl. Probab. 34, 702 (2002).
- [14] P. Calka, Adv. in Appl. Probab. 35, 551; 863 (2003).
- [15] A. Hayen and M. Quine, Adv. in Appl. Probab. 32, 67 (2000).
- [16] H.J. Hilhorst, in preparation.